

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
OAKLAND DIVISION

J. DOE 1, et al.,

Plaintiff(s),

vs.

GITHUB, INC., et al.,

Defendant(s).

~~Case Number: C-xx-xxxx~~

Case Numbers: 4:22-cv-06823-JST

4:22-cv-07074-JST

~~[MODEL] STIPULATION & PROPOSED~~
ORDER RE: DISCOVERY OF
ELECTRONICALLY STORED
INFORMATION ~~FOR PATENT~~
LITIGATION

Upon the stipulation of the parties, the Court ORDERS as follows:

1. This Order supplements all other discovery rules and orders. It streamlines Electronically Stored Information (“ESI”) production to promote a “just, speedy, and inexpensive determination of this action, as required by Federal Rule of Civil Procedure 1.”

2. This Order may be modified in the Court’s discretion or by stipulation. ~~The parties shall jointly submit any proposed modifications within 30 days after the Federal Rule of Civil Procedure 16 Conference.~~

3. As in all cases, costs may be shifted for disproportionate ESI production requests pursuant to Federal Rule of Civil Procedure 26. Likewise, a party’s nonresponsive or dilatory discovery tactics are cost-shifting considerations.

4. A party’s meaningful compliance with this Order and efforts to promote efficiency and reduce costs will be considered in cost-shifting determinations.

5. The parties are expected to comply with the District’s E-Discovery Guidelines (“Guidelines”) and are encouraged to employ the District’s Model Stipulated Order Re: the

1 Discovery of Electronically Stored Information and Checklist for Rule 26(f) Meet and Confer
2 regarding Electronically Stored Information.

3 6. General ESI production requests under Federal Rules of Civil Procedure 34 and 45
4 shall not include email or other forms of electronic correspondence or messaging service
5 (collectively “email”). To obtain email, parties must propound specific email production
6 requests.

7 7. Email production requests shall only be propounded for specific issues, rather than
8 general discovery of a product or business.

9 ~~8. Email production requests shall be phased to occur after the parties have exchanged~~
10 ~~initial disclosures and basic documentation about the patents, the prior art, the accused~~
11 ~~instrumentalities, and the relevant finances. While this provision does not require the production~~
12 ~~of such information, the Court encourages prompt and early production of this information to~~
13 ~~promote efficient and economical streamlining of the case.~~

14 98. Email production requests shall identify the custodian, search terms, and time frame.
15 The parties shall cooperate to identify the proper custodians, proper search terms and proper
16 timeframe as set forth in the Guidelines.

17 9. The parties will disclose at the outset any use of enterprise messaging tools, such as
18 Slack or MS Teams, including the type of subscription and retention settings in place. The
19 parties will meet and confer regarding issues relating to the parameters for collection of such
20 data, the format of the production, and other unique issues to this type of data.

21 10. Each requesting party shall limit its email production requests to a total of five
22 custodians per producing party for all such requests. OpenAI, Inc., OpenAI, L.P., OpenAI
23 OpCo, L.L.C., OpenAI GP, L.L.C., OpenAI Startup Fund GP I, L.L.C., OpenAI Startup Fund I,
24 L.P., and OpenAI Startup Fund Management, LLC shall be considered a single “producing
25 party.” The parties may jointly agree to modify this limit without the Court’s leave. The Court
26 shall consider contested requests for additional custodians, upon showing a distinct need based
27 on the size, complexity, and issues of this specific case. Cost-shifting may be considered as part
28 of any such request.

1 11. Each requesting party shall limit its email production requests to a total of five search
2 terms per custodian per producing party. All search terms per custodian per producing party shall
3 be identified in email production requests upon service. The parties may jointly agree to modify
4 this limit without the Court's leave. The Court shall consider contested requests for additional
5 search terms per custodian, upon showing a distinct need based on the size, complexity, and
6 issues of this specific case. The Court encourages the parties to confer on a process to test the
7 efficacy of the search terms. The search terms shall be narrowly tailored to particular issues.
8 Indiscriminate terms, such as the producing company's name or its product name, are
9 inappropriate unless combined with narrowing search criteria that sufficiently reduce the risk of
10 overproduction. A conjunctive combination of multiple words or phrases (*e.g.*, "computer" and
11 "system") narrows the search and shall count as a single search term. A disjunctive combination
12 of multiple words or phrases (*e.g.*, "computer" or "system") broadens the search, and thus each
13 word or phrase shall count as a separate search term unless they are variants of the same word.
14 Use of narrowing search criteria (*e.g.*, "and," "but not," "w/x") is encouraged to limit the
15 production and shall be considered when determining whether to shift costs for disproportionate
16 discovery. Should a party serve email production requests with search terms beyond the limits
17 agreed to by the parties or granted by the Court pursuant to this paragraph, this shall be
18 considered in determining whether any party shall bear all reasonable costs caused by such
19 additional discovery.

20 12. Nothing in this Order prevents the parties from agreeing to use technology assisted
21 review and other techniques insofar as their use improves the efficacy of discovery. Such topics
22 should be discussed pursuant to the District's E-Discovery Guidelines.

23 13. Production Format for ESI.

24 a. Except as otherwise provided in this Order, all documents existing in
25 electronic format (e.g., Word, PDF, webpage files) shall be produced as 300
26 DPI TIFF black and white images with a Bates number stamped on each
27 page, document metadata (including extracted or OCR text), and a link to
28 the associated extracted or OCR text file. The parties will consider

re-producing select and specifically identified Word documents as a color PDF upon reasonable request by the receiving party.

b. Source code and certain other highly confidential technical materials will be produced in native format and shall be made available for inspection on a secure computer in a secured room without Internet access or network access to other computers at a secure location determined by the producing party, pursuant to the terms of the Amended Protective Order (ECF No. 145) governing “HIGHLY CONFIDENTIAL – SOURCE CODE” materials.

c. Spreadsheet-type files (e.g., Microsoft Excel) and PowerPoint presentations should be produced in native format unless they require redaction. Documents produced in native format should be named according to the Bates number assigned, with the Confidentiality designation appended to the Bates numbers with a hyphen. Single-page Bates-stamped TIFF image slip-sheets will be included for each document produced in native format. The slip-sheets will display the Bates number of the native file, the Confidentiality endorsement, and an endorsement stating, “File Produced Natively.”

d. All attachments, addendums, enclosures, and/or exhibits to a parent document will be, to the extent reasonably possible, produced immediately following the parent document and identified as they relate to the respective parent document with the parent/child relationship intact.

e. To the extent reasonably possible, ESI items shall be processed so as to preserve the date/time shown in the document as it was last saved, not the date of collection or processing.

f. Telemetry data, training data, and other data inaccessible within the meaning of Federal Rule of Civil Procedure 26, shall not be produced to another party, either in native form or otherwise. The parties shall meet and

1 confer in good faith regarding reasonable methodologies to query or sample
2 such data sources. Given the large volume and nature of this data, searching
3 the entirety or even portions of such data is impossible, complex, limited,
4 and/or costly.

5 14. Metadata. Load or DAT files should include, where applicable and to the extent it
6 exists and was captured at the time of the collection, the information listed in **Exhibit A**
7 attached. To the extent that metadata does not exist or is not reasonably accessible or available
8 for any documents produced, nothing in this Order shall require any party to extract, capture,
9 collect or produce such metadata.

10 15. Privilege Logs. A party that withholds documents, ESI, and any other materials and
11 information on the grounds of attorney-client privilege, attorney work product, and/or any other
12 applicable privilege, immunity or protection from discovery, shall produce a privilege log.
13 Privileged communications exclusively between a party and its outside or in-house counsel on or
14 after November 3, 2022, and work product performed by a party at the direction of its outside or
15 in-house counsel on or after November 3, 2022 need not be included on a privilege log.
16 Privilege logs shall be served no later than 60 days before the close of fact discovery. The
17 privilege log shall include:

- 18 a. A unique number for each entry on the log.
19 b. Date of document for all ESI and to the extent known or discernible for all
20 hard copy documents. For emails this should be the sent date of the
21 document and for loose ESI this should be the last-modified or create date
22 of the document.
23 c. Custodian of the document. For emails this should be populated with the
24 metadata extracted from the Custodian(s) field. For loose ESI, this should
25 be populated with the Custodian value assigned to the instance of the
26 document in question.
27 d. Author of the document. For emails this should be populated with the
28 metadata extracted from the “Email From” field associated with the file.

For loose ESI, this should be populated with the metadata extracted from the “Author” field. For hard copy documents, this will be populated with the “Custodian.”

- e. Recipient(s) of the document where reasonably ascertainable. For emails this will be populated with the metadata extracted from the “Email To” field associated with the file. Separate columns will be included for the metadata extracted from the “Email CC” and “Email BCC” fields, where populated.
- f. The type of privilege being asserted, for example: (a) AC for Attorney/Client, (b) WP for Attorney Work Product, (c) CI for Common Interest.
- g. Nothing in Paragraph 15 of this Order shall be read to preclude the parties from logging documents categorically where appropriate.

16. Any Party that issues a non-party subpoena shall timely notify other Parties when it receives non-party productions of documents, ESI or other responses from the non-party, and shall provide copies of all such productions in the format in which they were received from the non-party to the other Parties. If the non-party produces documents or ESI that are not Bates labeled, the Receiving Party will Bates-label the documents or ESI before producing a copy to the other Parties if the produced documents or ESI are of the type so requiring consistent with other provisions of this Order. Nothing in this Protocol precludes a Receiving Party from reviewing a non-party production of documents or ESI and applying (if the non-party produced documents without a confidentiality designation) or elevating a confidentiality designation (e.g., elevating documents designated as CONFIDENTIAL to ATTORNEYS’ EYES ONLY) if the production contains confidential or extremely sensitive information as defined by the Amended Protective Order (ECF No. 145).

~~IT IS SO STIPULATED, through Counsel of Record.~~

Dated: _____

Counsel for Plaintiff

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

Dated: _____

Counsel for Defendant

IT IS ORDERED that the forgoing ~~Agreement~~ is approved.

Dated: _____

~~UNITED STATES DISTRICT/MAGISTRATE~~
~~JUDGE~~ Honorable Jon S. Tigar
United States District Judge

EXHIBIT A

<u>Field</u>	<u>Data Type</u>	<u>Paper</u>	<u>Native Files & Email Attachments</u>	<u>Email</u>
<u>BegDoc</u>	<u>Integer - Text</u>	<u>Starting Bates #</u>	<u>Starting Bates #</u>	<u>Starting Bates #</u>
<u>EndDoc</u>	<u>Integer - Text</u>	<u>Ending Bates #</u>	<u>Ending Bates #</u>	<u>Ending Bates #</u>
<u>BegAttach</u>	<u>Integer - Text</u>	<u>Starting bates # of document family</u>	<u>Starting bates # of document family</u>	<u>Starting bates # of document family</u>
<u>EndAttach</u>	<u>Integer - Text</u>	<u>Ending bates # of document family</u>	<u>Ending bates # of document family</u>	<u>Ending bates # of document family</u>
<u>Custodians</u>	<u>Text – paragraph Separate entries with “,”</u>		<u>All names of people the document was collected from even if removed from production as a duplicate</u>	<u>All names of people the document was collected from even if removed from production as a duplicate</u>
<u>From</u>	<u>Text - paragraph</u>			<u>Sender of message</u>
<u>To</u>	<u>Text – paragraph Separate entries with “,”</u>			<u>Recipients of message</u>
<u>CC</u>	<u>Text – paragraph Separate entries with “,”</u>			<u>Copied recipients</u>
<u>BCC</u>	<u>Text – paragraph Separate entries with “,”</u>			<u>Blind copied recipients</u>
<u>Subject</u>	<u>Text - paragraph</u>			<u>Subject of message</u>
<u>DateSent</u>	<u>Date (mm/dd/yy yy)</u>			<u>Date message sent</u>
<u>DateReceived</u>	<u>Date (mm/dd/yy yy)</u>			<u>Date message received</u>

<u><i>Field</i></u>	<u><i>Data Type</i></u>	<u><i>Paper</i></u>	<u><i>Native Files & Email Attachments</i></u>	<u><i>Email</i></u>
<u>FileName</u>	<u>Text - paragraph</u>		<u>Name of original file including extension</u>	<u>Name of original file including extension</u>
<u>FileExtension</u>	<u>Text</u>		<u>Extension of original file</u>	<u>Extension of original file</u>
<u>DateCreated</u>	<u>Date/Time (mm/dd/yy yy)</u>		<u>Date file was created</u>	
<u>DateModified</u>	<u>Date/Time (mm/dd/yy yy)</u>		<u>Date file was last modified</u>	
<u>Author</u>	<u>Text - paragraph</u>		<u>Document author from metadata</u>	
<u>Confidentiality</u>	<u>Single choice</u>	<u>Designation branded onto the document</u>	<u>Designation branded onto the document</u>	<u>Designation branded onto the document</u>
<u>FullText</u>	<u>Link</u>	<u>Link to the extracted or OCR text file</u>	<u>Link to the extracted or OCR text file</u>	<u>Link to the extracted or OCR text file</u>
<u>DocLink</u>	<u>Link</u>	<u>Link to the native file</u>	<u>Link to the native file</u>	<u>Link to the native file</u>
<u>MD5Hash</u>	<u>Text</u>		<u>The MD5Hash value of the file</u>	<u>The MD5Hash value of the file</u>